

Graph Decompositions at the Integral Expectation Threshold

OpenAI

October 5, 2026

Abstract

We prove the graph-decomposition conjecture of Ascoli, He, Park, and Talagrand. Every graph admits a partition into a universally bounded number of fixed edge pieces, each having ordinary containment threshold at most a universal constant times the original graph's integral expectation threshold. The partition is chosen before sampling the random host, and the separate embeddings of the pieces need not agree on shared vertices.

1 Introduction

All graphs are finite and simple, and copies are not required to be induced. For a graph J on at most n vertices, let

$$p_c(J; n) = \min\{p \in [0, 1] : \mathbb{P}(G(n, p) \text{ contains a copy of } J) \geq 1/2\}.$$

Here $G(n, p)$ has vertex set $[n] = \{1, \dots, n\}$, with edges present independently with probability p . We set $p_c(J; n) = 0$ when J is edgeless.

Remark (notation). G denotes a labelled graph on $[n]$, identified with its edge set in $2^{E(K_n)}$. In Proposition 3.2, a result cited in the proof of Proposition 3.2: In the containment configuration and the Hall argument, G is a fixed witness with $G \in \mathcal{D}$ and $E(J) \subseteq E(G)$. $p \in [0, 1]$ is the parameter used for edge inclusion and for evaluating cover cost. In Lemma 2.3: In the proof of the second bound, $p = 2^{-1/a} n^{-v/a} \in (0, 1)$, where the chosen subgraph has v vertices and $a \geq dv/2$ edges. In Proposition 4.2: In the remaining nontrivial case of the proof, $q_J = q(J; n) > 0$, $q_J < 1/(2A)$, and $p = Aq_J \in (0, 1/2)$, with $A = 400e^4$. $p_c(J; n)$ is the containment threshold for a finite simple graph J on at most n vertices, with copies not required to be induced and with $p_c(J; n) = 0$ for edgeless J .

To define the integral expectation threshold, put $X = E(K_n)$ and let $\mathcal{F}_H \subseteq 2^X$ be the increasing family of edge sets whose graphs contain H . A family $\mathcal{C} \subseteq 2^X$ covers $\mathcal{F}_H$ if every $A \in \mathcal{F}_H$ contains some $S \in \mathcal{C}$. Its cost at p is

$$c_p(\mathcal{C}) = \sum_{S \in \mathcal{C}} p^{|S|}.$$

For H with at least one edge, define

$$q(H; n) = \max\{p \in [0, 1] : \mathcal{F}_H \text{ has a cover } \mathcal{C} \text{ with } c_p(\mathcal{C}) \leq 1/2\}. \quad (1.1)$$

The maximum exists because there are only finitely many families of subsets of X . Each cover gives a first-moment bound on containment, so $q(H; n) \leq p_c(H; n)$. For convenience set $q(H; n) = 0$ for edgeless graphs as well.

Remark (notation). X is a finite ground set. In Theorem 1.1, Corollary 2.2, Lemma 2.3, Proposition 4.2: For graph-containment covers, $X = E(K_n)$. In Theorem 2.1: X is any nonempty finite set. S denotes a subset of the ground set in the covering relation, cost sum, or satisfaction test where it occurs. In Lemma 3.1, Proposition 3.2: As an argument of satisfaction or f_Q , $S \subseteq Y$; S satisfies Q exactly when it contains some member of Q . $q(H; n)$ is the integral expectation threshold defined by the cover-cost maximum for graphs H with $|V(H)| \leq n$, with $q(H; n) = 0$ when H is edgeless. In Theorem 1.1: After the edgeless case, the bare symbol q abbreviates $q(H; n) > 0$; in the subsequent nontrivial branch, $q < 1/(2A)$. A is context-dependent notation for a member of the family being covered or for a positive absolute numerical constant. In Theorem 2.1, Corollary 2.2, Lemma 2.3: In the covering definition, A ranges over the family being covered and is a subset of its ground set. In Proposition 4.2, Theorem 1.1: In numerical expressions, $A = 400e^4$. $c > 0$ denotes an absolute constant independent of the varying parameters; different estimates may use different values of c .

Theorem 1.1. *There exist an absolute integer $k \geq 2$ and an absolute constant $L \geq 1$ such that, for every $n \geq 2$ and every graph H on at most n vertices, the edges of H admit a partition*

$$E(H) = E(H_1) \dot{\cup} \cdots \dot{\cup} E(H_k) \quad \text{with} \quad p_c(H_i; n) \leq Lq(H; n) \quad (1 \leq i \leq k).$$

The pieces may be empty and may share vertices. The partition depends only on H, n and is fixed before the random host is sampled.

Remark (notation). For $1 \leq i \leq k$, H_i is a possibly empty edge piece of H in the partition $E(H) = E(H_1) \dot{\cup} \cdots \dot{\cup} E(H_k)$; the pieces may share vertices, and their separate embeddings need not agree on shared vertices.

This proves Conjecture 8.1 of Ascoli, He, Park, and Talagrand [1]. Their formulation allows overlapping edge sets. Assigning each edge to one piece containing it gives a partition and cannot increase any piece's containment threshold. Thus the two formulations are equivalent.

History and significance

Kahn and Kalai [5] proposed a general comparison between ordinary thresholds and integral expectation thresholds with a logarithmic loss. Following the fractional version established by Frankston, Kahn, Narayanan, and Park [2], Park and Pham [7] proved the integral comparison. For graph containment it gives

$$p_c(H; n) \leq Cq(H; n) \log \max\{2, |E(H)|\}$$

with an absolute constant C . Theorem 1.1 removes this logarithmic loss after partitioning the target into a bounded number of pieces. The assertion concerns each fixed piece's ordinary threshold; their embeddings need not agree on shared vertices.

Remark (notation). C denotes an unspecified positive absolute constant in the bound where it occurs; its occurrences do not specify a common numerical value.

Ascoli, He, Park, and Talagrand introduced generalized thresholds based on unions of likely graphs, related them to Talagrand's discrete-convexity conjecture, and formulated the decomposition problem above [1]. They proved bounded-degeneracy cases and further cases with both degeneracy and maximum-degree restrictions [1, Theorems 1.12–1.13]. Their Section 8 isolates the bipartite edges between high- and low-degree vertices as a principal remaining difficulty.

We use the same-parameter integral covering theorem from *Talagrand's discrete-convexity conjecture* [6, Theorem 1.1], which proves the discrete conjecture formulated in [8, Conjecture 7.1].

It is stated precisely in Section 2; this is an essential companion input. All remaining graph-decomposition and embedding arguments are proved below.

Method

The covering theorem places one labelled copy of H inside the union of a bounded number of deterministic graphs from a class $\mathcal{D}$ having high probability at a density r comparable to $q(H; n)$. Assigning each edge to one of these graphs gives initial pieces. In the large-degeneracy case we isolate a small set Y containing the high-degree vertices and enlarge it until each outside vertex has few neighbors in Y . Subsequent splitting produces pieces inside Y , outside Y , and across the cut.

Remark (notation). $r \in (0, 1)$ is the product-measure parameter, also used as the source random-graph edge density. In Corollary 2.2: $q(H; n) < r < 1$. In Proposition 3.2, Proposition 4.1: r and s satisfy $s \in \mathbb{N}$, $s \leq 10 \log n$, and $r^s \geq n^{-1/100}$. In Theorem 1.1: In the large-degeneracy case of the proof, $r = 2q(H; n)$. $\mathcal{D} \subseteq 2^X$ is a fixed family of subsets of a finite ground set X with $\mu_r(\mathcal{D}) \geq 1 - 1/K_0$. In Corollary 2.2: $X = E(K_n)$, and $\mathcal{D}$ is any family of graphs on $[n]$ satisfying the stated mass bound; it need not be the particular family defined by source tests. In Proposition 3.2, a result cited in the proof of Proposition 3.2, Theorem 1.1: $X = E(K_n)$ and $\mathcal{D} = \mathcal{D}(n, r, s)$ consists of exactly those graphs G for which, for every $Y \subseteq [n]$ with $|Y| \leq n^{2/3}$ and every nonempty $\mathcal{Q} \subseteq \{T \subseteq Y : 1 \leq |T| \leq s\}$ with $|\mathcal{Q}| \leq \lceil 2r^{-s} \rceil$, one has $\frac{1}{n-|Y|} \sum_{w \in [n] \setminus Y} f_{\mathcal{Q}}(N_G(w) \cap Y) \leq 2x(\mathcal{Q})$. Here $x(\mathcal{Q})$ is the probability that a Bernoulli- τ subset of Y satisfies $\mathcal{Q}$, with $\tau = 1 - (1 - r)^2$. In Theorem 1.1: In the large-degeneracy case, the chosen family uses $r = 2q(H; n)$ and $s = \lceil 2d/M \rceil$, where $d = \text{deg}_H(H)$ and $M = 1000$. Y is a finite set on which the requirement subsets and their satisfaction tests are defined. In Theorem 1.1: In the large-degeneracy construction, start with $Y_0 = \{v \in V(H) : \text{deg}_H(v) > \sqrt{n}\}$ and repeatedly add an outside vertex with more than $2d$ neighbors in the current set until none remains; Y is the resulting set. It satisfies $|Y| \leq 4d\sqrt{n} \leq n^{2/3}$, and every $w \notin Y$ has $\text{deg}_H(w) \leq \sqrt{n}$ and $|N_H(w) \cap Y| \leq 2d$. In Proposition 3.2, a result cited in the proof of Proposition 3.2: $Y \subseteq [n]$ with $|Y| \leq n^{2/3}$ is the part held fixed, and candidate images for vertices outside it are restricted to $[n] \setminus Y$.

To embed a cut piece while fixing Y , each nonisolated outside vertex needs a distinct image adjacent to its prescribed set of at most s vertices in Y . Hall's criterion asks for enough candidate images for every collection of outside vertices. Membership in a typical class alone does not ensure this; the class must control unions of arbitrarily many such conditions.

Remark (notation). s is an integer size bound for requirements or neighborhoods, with its value and additional restrictions specified locally. In Lemma 3.1: $s \geq 1$ is an integer and every $T \in \mathcal{R}$ satisfies $1 \leq |T| \leq s$. In Proposition 3.2, Proposition 4.1, a result cited in the proof of Proposition 3.2: $s \in \mathbb{N}$, $s \leq 10 \log n$, and $r^s \geq n^{-1/100}$, with $0 < r < 1$. In Proposition 3.2: s bounds the positive degrees in J of vertices outside Y . In Proposition 4.1: s bounds $\text{deg}_H(J)$. In Theorem 1.1: In the branch $d \geq M$, set $r = 2q$ and $s = \lceil 2d/M \rceil$, where $d = \text{deg}_H(H)$ and $M = 1000$.

The compression lemma selects at most $\lceil 2r^{-s} \rceil$ requirements from any family so that every neighborhood containing an original requirement contains a selected requirement with probability at least $1/2$ after adding each vertex of Y independently with probability r . This principle is independent of graphs. It permits simultaneous concentration over all subfamilies of this bounded size while still controlling every Hall condition. We define $\mathcal{D}$ using these tests for all small Y at once, so that transfer applies to cut pieces chosen after the deterministic cover.

For the two other regions, a direct embedding lemma uses successive matchings of layers of pairwise nonadjacent vertices. Graphs of bounded degeneracy are handled by small star forests, adapting the degeneracy splitting and center-testing arguments of [1, Claim 3.9 and Section 6].

Section 2 collects the covering input and elementary estimates. Section 3 proves neighborhood transfer, Section 4 gives the direct embedding lemmas, and Section 5 constructs the partition and proves Theorem 1.1.

2 Covering input and elementary estimates

The first step converts a family of typical labelled graphs into a deterministic cover of one copy of the target. We then record the elementary degree and threshold bounds used to subdivide that copy.

For a finite ground set X , write μ_r for the law of its Bernoulli- r random subset. Covering and the cost c_r have the same meaning as in the introduction, with X in place of $E(K_n)$.

Remark (notation). μ_r denotes the product law on subsets of the finite ground set in use, with each element included independently with probability r . In Corollary 2.2, Proposition 3.2: For a family $\mathcal{D}$ of graphs on $[n]$, $\mu_r(\mathcal{D})$ is evaluated on the ground set $E(K_n)$. In Lemma 3.1, Proposition 3.2, a result cited in the proof of Proposition 3.2: In the definition of f_Q , the ground set for $Z \sim \mu_r$ is Y .

Theorem 2.1 (Discrete convexity, [6, Theorem 1.1]). *Set $K_0 = 2^{75}$. For every nonempty finite set X , every $r \in (0, 1)$, and every family $\mathcal{D} \subseteq 2^X$ with $\mu_r(\mathcal{D}) \geq 1 - 1/K_0$, the family*

$$\mathcal{E}_{K_0}(\mathcal{D}) = \{S \subseteq X : S \not\subseteq D_1 \cup \dots \cup D_{K_0} \text{ for every } D_1, \dots, D_{K_0} \in \mathcal{D}\}$$

has a cover of cost at most $1/2$ at the same parameter r . The family $\mathcal{D}$ need not be monotone, and members of a tuple may repeat.

Remark (notation). $\mathcal{E}_{K_0}(\mathcal{D}) = \{S \subseteq X : S \not\subseteq D_1 \cup \dots \cup D_{K_0} \text{ for every } (D_1, \dots, D_{K_0}) \in \mathcal{D}^{K_0}\}$, where tuple entries may repeat and $\mathcal{D}$ need not be monotone.

Corollary 2.2. *Let $n \geq 2$, let H have an edge and at most n vertices, and suppose $q(H; n) < r < 1$. If a family of graphs on $[n]$ satisfies $\mu_r(\mathcal{D}) \geq 1 - 1/K_0$, then some copy of H has its edges contained in the union of K_0 members of $\mathcal{D}$.*

Remark. The members $D_1, \dots, D_{K_0} \in \mathcal{D}$ in the union $D_1 \cup \dots \cup D_{K_0}$ may repeat.

Proof. If no such copy exists, no graph containing H can be contained in such a union. Thus $\mathcal{F}_H \subseteq \mathcal{E}_{K_0}(\mathcal{D})$. Theorem 2.1 supplies a cover of $\mathcal{F}_H$ of cost at most $1/2$ at r , contrary to $r > q(H; n)$. $\square$

For a graph J , let

$$V_+(J) = \{v \in V(J) : \deg_J(v) > 0\}$$

be its set of nonisolated vertices. Isolated vertices can be deleted during an embedding and added afterward: any embedding of the remaining vertices extends injectively when $|V(J)| \leq n$. If $J \subseteq H$, then $\mathcal{F}_H \subseteq \mathcal{F}_J$, so every cover of $\mathcal{F}_J$ covers $\mathcal{F}_H$. In particular,

$$q(J; n) \leq q(H; n), \quad p_c(J; n) \leq p_c(H; n). \quad (2.1)$$

Remark (notation). v is locally bound, denoting either a vertex of the graph under discussion or the number of vertices of a locally chosen subgraph. In Lemma 2.3: In the proof of the second bound, v is the number of vertices of the chosen subgraph of H of minimum degree d , whose edge count is $a \geq dv/2$; in particular, $1 \leq v \leq n$.

The *degeneracy* $d = \text{degen}(H)$ is the maximum of the minimum degrees of nonempty subgraphs of H . Equivalently, the vertices admit an order with at most d preceding neighbors per vertex. To obtain the order, repeatedly remove a vertex of degree at most d and place it last. Conversely, the last vertex of any subgraph in such an order has degree at most d . Counting each edge at its later endpoint shows that every subgraph J of H satisfies

$$|E(J)| \leq d|V(J)|. \quad (2.2)$$

Remark (notation). d is a nonnegative integer degeneracy parameter. In Theorem 1.1, Lemma 2.3: For H having an edge, $d = \text{degen}(H) \geq 1$. In Lemma 2.4: The input graph has degeneracy at most d ; equality is not required.

Lemma 2.3. *If H has an edge, $|V(H)| \leq n$, and $d = \text{degen}(H)$, then*

$$q(H; n) \geq \frac{1}{2^{\binom{n}{2}}}, \quad q(H; n) \geq \frac{1}{2} n^{-2/d}. \quad (2.3)$$

Proof. The family of all single edges covers $\mathcal{F}_H$, proving the first bound. For the second, choose a subgraph of minimum degree d , with v vertices and $a \geq dv/2$ edges. Its copies on $[n]$ form a cover of $\mathcal{F}_H$ with cost at most $n^v p^a$. At $p = 2^{-1/a} n^{-v/a} \in (0, 1)$ this is at most $1/2$. Since $a \geq 1$ and $v/a \leq 2/d$, this parameter is at least $\frac{1}{2} n^{-2/d}$. $\square$

Remark (notation). a is a locally bound integer whose different occurrences in separate proofs have independent meanings. In Lemma 2.3: a is the number of edges of the chosen subgraph of minimum degree d , with v vertices, so $a \geq dv/2$ and $a \geq 1$. In Proposition 3.2: $a \in [n] \setminus Y$ is a candidate image for an outside vertex w , usable when $N_J(w) \subseteq N_{G_*}(a) \cap Y$. In Theorem 1.1: $a = |Y \setminus Y_0|$ counts the vertices added to Y_0 during the closure procedure.

We use the following elementary splitting from [1, Claim 3.9].

Lemma 2.4 (Degeneracy splitting). *For every positive integer m , the edges of a d -degenerate graph can be partitioned into m graphs, each of degeneracy at most $\lceil d/m \rceil$.*

Remark (notation). m is a locally specified positive integer. In Lemma 2.4: m is any positive integer specifying the number of graphs in the edge partition. In Proposition 3.2: For each test set Y , $m = n - |Y|$ is the number of vertices outside Y . In Proposition 4.2: For an integer $\ell \geq 1$ with $m_\ell > 0$, $m = m_\ell$ is the number of components with exactly ℓ leaves. In Theorem 1.1: In the applications of degeneracy splitting, $m = M = 1000$.

Proof. Fix a degeneracy order. At each vertex, partition its edges to preceding neighbors into m groups of size at most $\lceil d/m \rceil$, and assign each group to its corresponding piece. Each edge is assigned once. The same order certifies the claimed bound in every piece. $\square$

3 Uniform neighborhood transfer

We construct a typical class whose small-neighborhood bipartite subgraphs embed in a slightly denser random graph. To embed a bipartite graph J while fixing one part Y , each nonisolated

vertex w outside Y needs a distinct image adjacent to every vertex of $T_w = N_J(w) \subseteq Y$. For every set U of vertices to be matched, Hall's criterion asks for at least $|U|$ candidate images whose neighborhoods contain at least one T_w with $w \in U$. The family of requirements indexed by U can be large. The next lemma reduces it to a bounded subfamily after an independent random extension.

Remark (notation). T is local notation whose domain and meaning are determined by its occurrence. In Lemma 3.1: $T \subseteq Y$ is a requirement set with $1 \leq |T| \leq s$; in the selection procedure and its termination argument, $T \in \mathcal{R}$. In Proposition 3.2: T is a subset of Y with $1 \leq |T| \leq s$, indexing the allowed members of a test family. In Proposition 4.2: T is the random total number of center trials needed to finish all stars after infinitely many auxiliary centers with fresh independent rows have been appended; trials include failures. U indexes vertices to be matched to distinct candidate images. In Proposition 3.2: U is any nonempty subset of the positive-degree vertices of J in $[n] \setminus Y$, and its requirements are $\mathcal{R} = \{N_J(w) : w \in U\}$. In Proposition 4.1: U is any nonempty subset of the current layer, with $|U| = u$; its requirements are adjacency to the already embedded neighbors in earlier layers, and its candidate images lie in the current bin. $N_J(w) = \{v \in V(J) : \{v, w\} \in E(J)\}$ for $w \in V(J)$. In Proposition 3.2: For a positive-degree vertex $w \in [n] \setminus Y$ to be matched, $N_J(w) \subseteq Y$ and $1 \leq |N_J(w)| \leq s$.

3.1 Compressing requirements

For a family $\mathcal{Q}$ of subsets of a finite set Y , say that $S \subseteq Y$ *satisfies* $\mathcal{Q}$ if it contains some member of $\mathcal{Q}$. Given $r \in (0, 1)$, define

$$f_{\mathcal{Q}}(S) = \mathbb{P}(S \cup Z \text{ satisfies } \mathcal{Q}), \quad Z \sim \mu_r \text{ on } Y. \quad (3.1)$$

A small subfamily need not be satisfied by every set that satisfies the original family. The lemma replaces this implication by a uniform probability bound after the extension.

Remark (notation). $\mathcal{Q} \subseteq 2^Y$ is a family of requirements, and a set $S \subseteq Y$ satisfies it exactly when some $T \in \mathcal{Q}$ is contained in S . In Lemma 3.1: The asserted $\mathcal{Q}$ is a nonempty subfamily of $\mathcal{R} \subseteq \{T \subseteq Y : 1 \leq |T| \leq s\}$ with $|\mathcal{Q}| \leq \lceil 2r^{-s} \rceil$, such that every S satisfying $\mathcal{R}$ has $f_{\mathcal{Q}}(S) \geq 1/2$. During its construction the proof also uses $\mathcal{Q}$ for the growing family initialized at $\emptyset$. In Proposition 3.2, a result cited in the proof of Proposition 3.2: The simultaneous tests range over every nonempty $\mathcal{Q} \subseteq \{T \subseteq Y : 1 \leq |T| \leq s\}$ with $|\mathcal{Q}| \leq \lceil 2r^{-s} \rceil$. In Proposition 3.2: For a Hall test U , choose $\mathcal{Q} \subseteq \mathcal{R} = \{N_J(w) : w \in U\}$ by compression; this choice may depend on J, U, G , since the source and target inequalities hold simultaneously for every admissible test.

Lemma 3.1 (Compression after a random extension). *Let Y be finite, $r \in (0, 1)$, and $s \geq 1$ an integer. For every nonempty family*

$$\mathcal{R} \subseteq \{T \subseteq Y : 1 \leq |T| \leq s\}$$

there is a nonempty subfamily $\mathcal{Q} \subseteq \mathcal{R}$ with $|\mathcal{Q}| \leq \lceil 2r^{-s} \rceil$ such that

$$S \text{ satisfies } \mathcal{R} \implies f_{\mathcal{Q}}(S) \geq 1/2. \quad (3.2)$$

Remark (notation). $\mathcal{R}$ is a nonempty family of subsets of Y , each of size between 1 and s . In Proposition 3.2: For a nonempty set U of vertices to be matched, $\mathcal{R} = \{N_J(w) : w \in U\}$, with repeated neighborhoods represented only once.

Proof. Start with $\mathcal{Q} = \emptyset$. While some $T \in \mathcal{R}$ satisfies

$$\mathbb{P}(Z \text{ satisfies } \mathcal{Q} \mid Z \supseteq T) < 1/2,$$

add T to $\mathcal{Q}$. The probability that Z satisfies the current family increases by

$$\mathbb{P}(Z \supseteq T, Z \text{ does not satisfy } \mathcal{Q}) > \frac{1}{2}r^{|T|} \geq \frac{1}{2}r^s$$

at each addition. A previously added set cannot be chosen again, and the process makes fewer than $2r^{-s}$ additions. It makes at least one, because $\mathcal{R}$ is nonempty.

At termination the conditional probability is at least $1/2$ for every $T \in \mathcal{R}$. The conditional law of Z given $Z \supseteq T$ is the law of $T \cup Z$ with a fresh Bernoulli sample. Hence $f_{\mathcal{Q}}(T) \geq 1/2$. The function $f_{\mathcal{Q}}$ is increasing in S , proving (3.2) whenever S contains such a T . $\square$

3.2 The transfer lemma

All logarithms below are natural. The parameter range used in both this section and the next is

$$0 < r < 1, \quad s \in \mathbb{N}, \quad s \leq 10 \log n, \quad r^s \geq n^{-1/100}. \quad (3.3)$$

The bounds will be uniform over this range.

Proposition 3.2 (Neighborhood transfer). *There is an absolute integer N_1 such that the following holds for every $n \geq N_1$ and every r, s satisfying (3.3). There exists $\mathcal{D} = \mathcal{D}(n, r, s) \subseteq 2^{E(K_n)}$ with $\mu_r(\mathcal{D}) \geq 1 - 1/K_0$ such that, whenever*

- $Y \subseteq [n]$ has $|Y| \leq n^{2/3}$;
- J is a labelled bipartite graph with all edges between Y and $[n] \setminus Y$, and $E(J) \subseteq E(G)$ for some $G \in \mathcal{D}$;
- at most $n/8$ vertices outside Y have positive degree in J , and all of those degrees are at most s ;

one has

$$p_c(J; n) \leq r_* := 1 - (1 - r)^{32} \leq 32r. \quad (3.4)$$

Proof. We impose upper bounds on source sums of $f_{\mathcal{Q}}$ and lower bounds on target counts satisfying $\mathcal{Q}$. Compression makes each vertex in a set U to be matched contribute at least $1/2$ to an appropriate source sum, while each target neighborhood counted by the corresponding $\mathcal{Q}$ is usable for at least one vertex of U .

The test families. Use every $Y \subseteq [n]$ of size at most $n^{2/3}$ and every nonempty family

$$\mathcal{Q} \subseteq \{T \subseteq Y : 1 \leq |T| \leq s\}, \quad |\mathcal{Q}| \leq \lceil 2r^{-s} \rceil.$$

Write $m = n - |Y|$, put $\tau = 1 - (1 - r)^2$, and let $x(\mathcal{Q})$ and $x_*(\mathcal{Q})$ be the probabilities that Bernoulli- τ and Bernoulli- r_* subsets of Y satisfy $\mathcal{Q}$, respectively. Both are at least r^s . Define $\mathcal{D}$ by requiring, for every test $(Y, \mathcal{Q})$,

$$\frac{1}{m} \sum_{w \in [n] \setminus Y} f_{\mathcal{Q}}(N_G(w) \cap Y) \leq 2x(\mathcal{Q}). \quad (3.5)$$

Also define a target event by the inequalities

$$\frac{1}{m} |\{w \in [n] \setminus Y : N_{G_*}(w) \cap Y \text{ satisfies } \mathcal{Q}\}| \geq \frac{3}{4} x_*(\mathcal{Q}) \quad (3.6)$$

for all of the same tests.

Remark (notation). G_* denotes the target graph on $[n]$ at density r_* . In Proposition 3.2: The target-event probability is evaluated under $G_* \sim G(n, r_*)$. In the Hall argument, G_* is any fixed realization satisfying (3.6) for every admissible test, and Y is kept fixed during the embedding.

Simultaneous concentration. For fixed $Y, \mathcal{Q}$ and $G \sim G(n, r)$, the summands in (3.5) are independent and lie in $[0, 1]$. Their mean is $x(\mathcal{Q})$: the neighborhood sample and the independent sample in (3.1) have union density τ . Their independence follows because the edges from distinct vertices outside Y into Y are disjoint sets of random variables.

The logarithm of the number of tests is at most

$$C(n^{2/3} \log n + n^{1/100} (\log n)^2) \quad (3.7)$$

with an absolute C . Indeed, choosing Y costs at most $(\lfloor n^{2/3} \rfloor + 1)n^{n^{2/3}}$ possibilities. One can then list at most $\lceil 2n^{1/100} \rceil$ subsets, each of size at most $s \leq 10 \log n$. This overcounts families and gives (3.7).

For large n , $m \geq n/2$. Since $x(\mathcal{Q}) \geq n^{-1/100}$, Hoeffding's inequality [4] bounds the failure probability of (3.5) for a fixed test by $\exp(-cn^{98/100})$. The indicators in (3.6), for $G_* \sim G(n, r_*)$, are likewise independent, with mean $x_*(\mathcal{Q}) \geq n^{-1/100}$. The same bound, with another absolute $c > 0$, applies to their relative deviation by $1/4$. Thus each simultaneous event has probability at least $1 - \varepsilon_n$, where uniformly in r, s ,

$$\varepsilon_n \leq 2 \exp\left(C(n^{2/3} \log n + n^{1/100} (\log n)^2) - cn^{98/100}\right) \longrightarrow 0. \quad (3.8)$$

Increase N_1 so that this error is at most $\min\{1/K_0, 1/2\}$ for $n \geq N_1$.

The relation between the two test means is important. A Bernoulli- r_* set is the union of 16 independent Bernoulli- τ sets. Satisfaction is increasing, so

$$x_*(\mathcal{Q}) \geq 1 - (1 - x(\mathcal{Q}))^{16}. \quad (3.9)$$

Hall's criterion. Fix J, Y, G as in the statement and a target graph G_* satisfying (3.6). Keep Y fixed and match the positive-degree vertices of J outside Y to distinct vertices of $[n] \setminus Y$. A candidate image a is usable for a vertex w when $N_J(w) \subseteq N_{G_*}(a) \cap Y$.

For any nonempty set U of vertices to be matched, apply Lemma 3.1 to

$$\mathcal{R} = \{N_J(w) : w \in U\}.$$

It gives an admissible test $\mathcal{Q} \subseteq \mathcal{R}$. Every $N_G(w) \cap Y$, for $w \in U$, satisfies $\mathcal{R}$, because G contains the labelled edges of J . Writing $x = x(\mathcal{Q})$, (3.5) therefore yields

$$|U|/2 \leq \sum_{w \in [n] \setminus Y} f_{\mathcal{Q}}(N_G(w) \cap Y) \leq 2xm. \quad (3.10)$$

Repeated neighborhoods disappear from $\mathcal{R}$, but their distinct vertices still contribute separately to this sum.

Every target vertex counted in (3.6) is usable for some vertex of U , because $\mathcal{Q} \subseteq \mathcal{R}$. If $x \leq 1/16$, then

$$x_*(\mathcal{Q}) \geq 1 - e^{-16x} \geq 8x,$$

so there are at least $6xm \geq |U|$ usable vertices. If $x > 1/16$, there are at least

$$\frac{3}{4}(1 - e^{-1})m \geq \frac{3}{8}(1 - e^{-1})n > n/8 \geq |U|.$$

Hall's theorem [3] gives the matching in both cases. All choices of $\mathcal{Q}$ are legitimate even though they depend on J, U, G : the inequalities were established simultaneously for every test.

The matching embeds every edge of J while fixing Y . Any remaining isolated vertices can be added afterward. The target event has probability at least $1/2$, giving $p_c(J; n) \leq r_*$. Finally, the union bound for 32 independent Bernoulli- r samples gives $r_* \leq 32r$. $\square$

4 Direct embedding lemmas

Neighborhood transfer will handle edges between a small high-degree set and its complement. Two direct estimates cover the other parts: one for graphs with bounded maximum degree and one for small star forests. Neither requires membership in the typical family.

4.1 A maximum-degree bound

Proposition 4.1. *There is an absolute integer N_2 such that, for $n \geq N_2$ and r, s satisfying (3.3), every graph J with*

$$|V_+(J)| \leq n/4, \quad \text{deg}_n(J) \leq s, \quad \Delta(J) \leq n^{2/3}$$

and at most n vertices satisfies $p_c(J; n) \leq r$.

Proof. Discard isolated vertices and assume that an edge remains. We partition the vertices into layers that are independent sets, constructed in reverse order. In any residual subgraph with m' vertices, the average degree is at most $2s$. At least $m'/2$ vertices therefore have degree at most $4s$. The graph they induce has an independent set of size at least $m'/(2(4s+1))$, by greedily choosing a vertex and deleting it and its neighbors. Take such a set as the last remaining layer and repeat on the residual graph. Geometric decrease gives

$$t \leq Cs \log n \leq C'(\log n)^2 \tag{4.1}$$

layers. Every vertex has at most $4s$ neighbors in earlier layers.

Remark (notation). t is the number of independent-set layers constructed in reverse order after isolated vertices are discarded: each selected layer has size at least $m'/(2(4s+1))$ in the current residual graph on m' vertices, and every vertex has at most $4s$ neighbors in earlier layers.

For each layer choose a disjoint bin in $[n]$ of size equal to the layer size plus

$$b = \left\lceil \frac{n}{4t} \right\rceil.$$

The bins fit, since their total size is at most $n/4 + tb \leq n/2$. Embed the layers in order. At a given step expose the edges between the current bin and the earlier images, and seek a matching

of the layer into its bin. These edges are fresh independent Bernoulli- r variables conditional on the entire past: all previous exposures were between earlier bins.

Write $\Delta = \Delta(J)$. Fix a subset U of the current layer with $|U| = u \geq 1$. Each vertex of U requires adjacency to at most $4s$ earlier images. Each earlier image occurs in at most Δ requirements, so the conflict graph joining intersecting requirements has maximum degree at most $4s\Delta$. It has an independent set of size at least $u/(1 + 4s\Delta)$, whose requirements are pairwise disjoint.

Remark (notation). $\Delta(J)$ denotes the maximum degree of J . In Proposition 4.1: The bare symbol Δ abbreviates $\Delta(J)$ for the graph being embedded, with $\Delta(J) \leq n^{2/3}$. $u = |U| \geq 1$ is the cardinality of the nonempty set being tested by Hall's criterion. In Proposition 4.1: $1 \leq u \leq$ the size of the current layer; the corresponding set W consists of $u - 1$ vertices of its bin.

For a set W of $u - 1$ vertices in the current bin, at least b vertices of the bin remain outside W . The probability that none is usable for any vertex of U is at most

$$\exp\left(-\frac{bur^{4s}}{1 + 4s\Delta}\right). \quad (4.2)$$

Indeed, restrict to the disjoint requirements just selected. Each succeeds at a given candidate with probability at least r^{4s} , independently over those requirements and over candidates. If an empty requirement is present, the failure event is impossible, so the bound remains valid.

Uniformly under (3.3) and (4.1), for sufficiently large n ,

$$\frac{br^{4s}}{1 + 4s\Delta} \geq c \frac{n^{1-4/100-2/3}}{(\log n)^3} = c \frac{n^{22/75}}{(\log n)^3} \geq 5 \log n. \quad (4.3)$$

Here $b \geq c'n/(\log n)^2$ and $r^{4s} \geq n^{-4/100}$. There are at most n^{2u} choices of U, W for each u . Failure of Hall's condition implies such a pair with all usable images of U in W . By (4.2)–(4.3), the conditional failure probability at a layer is at most

$$\sum_{u \geq 1} n^{2u} e^{-5u \log n} = \sum_{u \geq 1} n^{-3u} = O(n^{-3}).$$

Summing this bound over the $O((\log n)^2)$ layers shows that the embedding succeeds with probability tending to one, uniformly over the allowed parameters. Taking an absolute N_2 so that it succeeds with probability at least $1/2$ proves the proposition. $\square$

4.2 Small star forests

A star forest is a graph whose nontrivial components are stars. The next argument adapts the fresh-center embedding in [1, Claim 6.6]. Small support is imposed directly, so it is unnecessary to divide the stars by their degrees in advance.

Proposition 4.2. *Put $A = 400e^4$. There is an absolute integer N_3 such that, for every $n \geq N_3$ and every star forest J on at most n vertices with $|V_+(J)| \leq n/100$,*

$$p_c(J; n) \leq 2Aq(J; n). \quad (4.4)$$

If $q(J; n) < 1/(2A)$, the stronger bound $p_c(J; n) \leq Aq(J; n)$ holds.

Proof. The edgeless case is immediate. Put $q_J = q(J; n) > 0$. If $q_J \geq 1/(2A)$, then $2Aq_J \geq 1$, so (4.4) is automatic. We may therefore assume $q_J < 1/(2A)$ and set $p = Aq_J \in (0, 1/2)$.

Remark (notation). $q_J = q(J; n)$. In Proposition 4.2: The abbreviation is introduced for J with at least one edge, so $q_J > 0$; in the remaining branch of the proof, $q_J < 1/(2A)$ and $p = Aq_J$.

For $\ell \geq 1$, let m_ℓ be the number of components of J with exactly ℓ leaves, ignoring isolated vertices. If $m = m_\ell > 0$, the copies on $[n]$ of m disjoint ℓ -edge stars cover $\mathcal{F}_J$. Their number is at most

$$C_\ell = \frac{n^{m(\ell+1)}}{(\ell!)^m m!}.$$

This follows by counting vertex injections and dividing by permutations of the leaves and of the stars. For $\ell = 1$ there are additional automorphisms, which only decrease the actual count. Since such a star forest exists on at most n vertices, $C_\ell \geq 1$. Using its cover at $(1/(2C_\ell))^{1/(m\ell)} \in (0, 1)$ yields

$$q_J^\ell \geq \frac{\ell!}{n^{\ell+1}} (m!/2)^{1/m} \geq \frac{\ell!}{n^{\ell+1}} \frac{m_\ell}{2e}. \quad (4.5)$$

The last step uses $m! \geq (m/e)^m$ and $2^{1/m} \leq 2$.

Remark (notation). $C_\ell = n^{m(\ell+1)} / ((\ell!)^m m!)$ is defined for integers $\ell \geq 1$ with $m = m_\ell > 0$ and bounds the number of copies on $[n]$ of m disjoint ℓ -edge stars.

Partition $[n]$ into a center bin of size $\lfloor n/2 \rfloor$ and a leaf bin containing the remaining vertices. Process the components of J one at a time. For a star with ℓ leaves, test unused centers successively until one has at least ℓ neighbors among the unused leaf vertices. Use that center and any ℓ such neighbors, and never test that center again. Failed centers are also discarded.

Throughout the process, at least $b' = \lfloor n/3 \rfloor$ leaf vertices remain. The unused leaf set is determined by previously tested center rows, and a fresh center row has not been exposed. Thus each trial, conditional on the past, has success probability at least

$$u_\ell = \mathbb{P}(\text{Bin}(b', p) \geq \ell). \quad (4.6)$$

For the purpose of counting trials, append infinitely many auxiliary centers with fresh independent rows to the same leaf bin. This enlarged process agrees with the actual process until the original center bin is exhausted. The number T of trials needed to finish all stars in the enlarged process satisfies

$$\mathbb{E}T \leq \sum_{\ell: m_\ell > 0} \frac{m_\ell}{u_\ell}, \quad (4.7)$$

because the conditional expected waiting time for each success is at most $1/u_\ell$.

Remark (notation). $u_\ell = \mathbb{P}(\text{Bin}(b', p) \geq \ell)$ is a binomial-tail lower bound on the success probability of each trial for an ℓ -leaf star conditional on the past.

If $\mu = b'p \geq 2\ell$, the second-moment bound $\mathbb{E}[\text{Bin}(b', p)^2] \leq \mu^2 + \mu$ and the Paley–Zygmund inequality give

$$u_\ell \geq \frac{1}{4} \frac{\mu^2}{\mu^2 + \mu} \geq \frac{1}{6},$$

since $\mu \geq 2$. Otherwise $b'p < 2\ell$. For every occupied star size, $\ell \leq n/100$, and for large n , $b' - \ell + 1 \geq n/4$. As $p < 1/2$, the inequality $\log(1 - p) \geq -2p$ implies

$$\begin{aligned} u_\ell &\geq \binom{b'}{\ell} p^\ell (1 - p)^{b' - \ell} \geq \frac{(n/4)^\ell p^\ell}{\ell!} e^{-4\ell} \\ &\geq \left(\frac{A}{4e^4} \right)^\ell \frac{m_\ell}{2en} = 100^\ell \frac{m_\ell}{2en}, \end{aligned} \quad (4.8)$$

where the last inequality uses (4.5).

There are at most $n/200$ nontrivial components. Combining the two bounds in (4.7) gives

$$\mathbb{E}T \leq 6\frac{n}{200} + 2en \sum_{\ell \geq 1} 100^{-\ell} = \left(\frac{3}{100} + \frac{2e}{99} \right) n < \frac{n}{8}.$$

For large n , Markov's inequality shows that $T \leq \lfloor n/2 \rfloor$ with probability at least $1/2$. On this event all centers used belong to the actual random graph, so the construction embeds J in $G(n, p)$. This proves $p_c(J; n) \leq Aq_J$ in the remaining case. $\square$

5 Constructing the fixed partition

We now prove Theorem 1.1. The high-degeneracy case first uses the typical-family cover and isolates a small set Y so that every vertex outside Y has few neighbors in Y and bounded total degree. Neighborhood transfer handles edges across this cut, while the maximum-degree estimate handles edges inside and outside Y after further splitting. The bounded-degeneracy case uses only small star forests. The partition choices in both cases are deterministic.

Proof of Theorem 1.1. Fix

$$M = B = 1000, \quad k = 2K_0 M(B+1)^2. \quad (5.1)$$

Choose an absolute integer $n_0 \geq 2$ large enough for Propositions 3.2, 4.1, and 4.2, and so that, for every $n \geq n_0$,

$$\frac{8 \log n}{\log 2} \sqrt{n} \leq n^{2/3}, \quad n^{2/3} \leq n/4, \quad 2\lceil n/B \rceil \leq n/100. \quad (5.2)$$

All these conditions have absolute cutoffs. Set

$$L = \max\{2A, 64, n_0^2\}.$$

If H is edgeless, use only empty pieces. Otherwise write $q = q(H; n) > 0$ and $d = \deg(H) \geq 1$. When $n < n_0$, Lemma 2.3 gives $Lq \geq n_0^2/(n(n-1)) \geq 1$. When $q \geq 1/(2A)$, we also have $Lq \geq 1$. In either case put all edges in one piece and use empty pieces for the rest. Henceforth

$$n \geq n_0, \quad q < 1/(2A). \quad (5.3)$$

Large degeneracy: $d \geq M$. Set

$$r = 2q, \quad s = \lceil 2d/M \rceil.$$

By Lemma 2.3, $r \geq n^{-2/d}$, whereas (5.3) gives $r < 1/A < 1/2$. Consequently

$$d < \frac{2 \log n}{\log 2}, \quad s \leq 3d/M, \quad r^s \geq n^{-6/M} \geq n^{-1/100}. \quad (5.4)$$

These bounds also give $s \leq 10 \log n$, so (3.3) holds.

Take the family $\mathcal{D}$ from Proposition 3.2. Corollary 2.2 supplies a copy of H covered by K_0 members $G_1, \dots, G_{K_0}$ of $\mathcal{D}$. Fix those graphs and the labelled copy, and use this copy to label

H on a subset of $[n]$. Assign each edge of H to one G_i containing it. This gives K_0 initial edge pieces, each contained in its specified typical graph.

We separate the high-degree vertices, following the obstacle identified in [1, Section 8]. Initially let

$$Y_0 = \{v \in V(H) : \deg_H(v) > \sqrt{n}\}.$$

By (2.2), $|Y_0| \leq 2d\sqrt{n}$. Starting from Y_0 , add an outside vertex whenever it has more than $2d$ neighbors in the current set, and stop when no such vertex remains. Let Y be the resulting set and let $a = |Y \setminus Y_0|$. Each added vertex contributes more than $2d$ edges to earlier vertices of the set; no edge is counted twice. Thus

$$2da \leq |E(H[Y])| \leq d(|Y_0| + a),$$

giving $|Y| \leq 2|Y_0| \leq 4d\sqrt{n} \leq n^{2/3}$ by (5.4) and (5.2). The conclusion also holds if Y_0 is empty, in which case no vertex is added. At termination every $w \notin Y$ satisfies

$$\deg_H(w) \leq \sqrt{n}, \quad |N_H(w) \cap Y| \leq 2d. \quad (5.5)$$

Partition $V(H) \setminus Y$ into B possibly empty blocks of size at most $\lceil n/B \rceil$. Subdivide each initial edge piece as follows.

Edges inside Y . Apply Lemma 2.4 with $m = M$ to obtain M pieces of degeneracy at most $\lceil d/M \rceil \leq s$. Each is supported on at most $|Y| \leq n^{2/3} \leq n/4$ vertices and has maximum degree at most $n^{2/3}$. Proposition 4.1 gives threshold at most r .

Edges outside Y . First partition by the unordered pair of endpoint blocks, allowing both endpoints in the same block. There are $B(B+1)/2$ such pairs. Apply Lemma 2.4 with $m = M$ to each resulting graph. The pieces have degeneracy at most s , maximum degree at most $\sqrt{n}$, and support size at most $2\lceil n/B \rceil \leq n/100$. Proposition 4.1 again gives threshold at most r .

Edges between Y and its complement. First partition by the block of the endpoint outside Y . At each such endpoint, distribute its edges in the initial piece among M groups of size at most $\lceil 2d/M \rceil = s$, using (5.5). This assigns every cut edge once. Each resulting bipartite piece is still contained in its specified $G_i \in \mathcal{D}$, has at most $\lceil n/B \rceil \leq n/8$ positive-degree vertices outside Y , and has all those degrees at most s . Proposition 3.2 gives threshold at most $r_* \leq 32r = 64q$.

All pieces therefore have threshold at most $64q \leq Lq$. Their number is at most

$$K_0 M \left(1 + \frac{B(B+1)}{2} + B \right) = \frac{K_0 M (B+1)(B+2)}{2} \leq k.$$

Bounded degeneracy: $1 \leq d < M$. Lemma 2.4, with $m = d$, partitions $E(H)$ into d graphs of degeneracy at most one, hence into forests. Indeed, the last vertex of a cycle in the certifying order would have two preceding neighbors.

Root every nontrivial tree. Color each parent–child edge by the parity of the depth of its parent. Each color class is a star forest: its centers occur at one depth parity, and their children have the other parity and no other incident edge of that color. This is the parity splitting used in [1, Claim 6.2].

Partition $V(H)$ into B blocks of size at most $\lceil n/B \rceil$, and split each star forest by unordered endpoint-block pairs. Every resulting piece J remains a star forest and has $|V_+(J)| \leq 2\lceil n/B \rceil \leq n/100$. By Proposition 4.2 and (2.1),

$$p_c(J; n) \leq 2Aq(J; n) \leq 2Aq(H; n) \leq Lq.$$

The number of pieces is at most $2dB(B+1)/2 = dB(B+1) \leq k$.

Finally, add empty pieces if necessary. In the large-degeneracy case, pull the labelled edge pieces back to the original H . The typical graphs, the copy of H , all vertex orders, and all assignments have been fixed using only H, n ; none depends on the random host used to test a piece's threshold. This completes the proof. $\square$

References

- [1] Ruben Ascoli, Xiaoyu He, Jinyoung Park, and Michel Talagrand, *A reformulation of the discrete Convexity Conjecture via k -thresholds*, arXiv:2608.11183v1, 2026, <https://arxiv.org/abs/2608.11183v1>.
- [2] Keith Frankston, Jeff Kahn, Bhargav Narayanan, and Jinyoung Park, *Thresholds versus fractional expectation-thresholds*, Annals of Mathematics **194** (2021), no. 2, 475–495.
- [3] Philip Hall, *On representatives of subsets*, Journal of the London Mathematical Society **10** (1935), no. 1, 26–30.
- [4] Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association **58** (1963), no. 301, 13–30.
- [5] Jeff Kahn and Gil Kalai, *Thresholds and expectation thresholds*, Combinatorics, Probability and Computing **16** (2007), no. 3, 495–502.
- [6] OpenAI, *Talagrand's discrete-convexity conjecture*, OpenAI Math Release preprint [OAI:Talagrands-discrete-convexity-conjecture-September-23-2026](https://arxiv.org/abs/2608.11183v1), 2026.
- [7] Jinyoung Park and Huy Tuan Pham, *A proof of the Kahn–Kalai conjecture*, Journal of the American Mathematical Society **37** (2024), no. 1, 235–243.
- [8] Michel Talagrand, *Are many small sets explicitly small?*, Proceedings of the Forty-Second ACM Symposium on Theory of Computing (New York), ACM, 2010, pp. 13–36.